

一种手机监控通信协议的设计与实现

Design and Implementation of a Communications Protocol for Mobile Phone Monitoring Systems

张纯容 张洪渊^① 王 忠 周庆标 (温州大学 城市学院网络系 浙江 温州 325035)

摘 要: 手机视频监控系统结合 GPRS 移动通信技术,把智能手机作为监控系统的显示和控制终端。本文详细介绍了一种智能手机和流媒体服务器之间的通信协议的设计和实现。通过该协议,可以在智能手机上查看远程视频监控画面并对监控摄像头进行左转、右转、上仰、下俯、设置预置位及转到给定预置位等操作。

关键词: 手机监控通信协议 PTZ 控制 远程视频监控 流媒体服务器 手机视频监控

1 引言

近年来视频监控越来越成为安防领域关注的热点,并已广泛应用于社会治安、营业场所、消防等各个领域。目前市场上的视频监控系统主要采用有线的通信方式,这种方式需要建立监控室,安排专人值守,当监控室无人或者没有可用的联网终端时,人们就无法及时查看到视频监控画面。在一定程度上削弱了监控系统的功能,降低了系统的使用效果。

手机视频监控系统则充分利用现有的网络视频监控技术,它结合 GPRS 移动通信技术,把智能手机作为监控系统的一种显示控制终端,从而通过智能手机终端实现远程监控画面的查看及监控摄像头的 PTZ (Pan - Tilt - Zoom 即左右转动,上下俯仰及调焦等功能)控制操作。由于手机视频监控系统的可移动性,它有效的提高了现有监控系统的使用效率,同时将监控系统的应用范围由目前的行业应用和安防应用领域拓宽到个人、家庭用户以及非安防应用领域。

在手机视频监控系统中,手机除了具备接收来自流媒体服务器的视频数据之外,还必须具备能够与流媒体服务器进行双向交互通信的功能,以便向流媒体服务器发出登陆和 PTZ 控制等信息。本文设计与实现的手机监控通信协议 MMCP (Mobile Phone Monitoring Communication Protocol),语法简洁、语意简化、网络通信的开销少,有效的解决了同步问题,实现了手机在查看视频监控画面的同时还可以与流媒体服务器通信,

并通过流媒体服务器控制摄像头 PTZ 动作的功能。

2 手机视频监控系统的组网结构

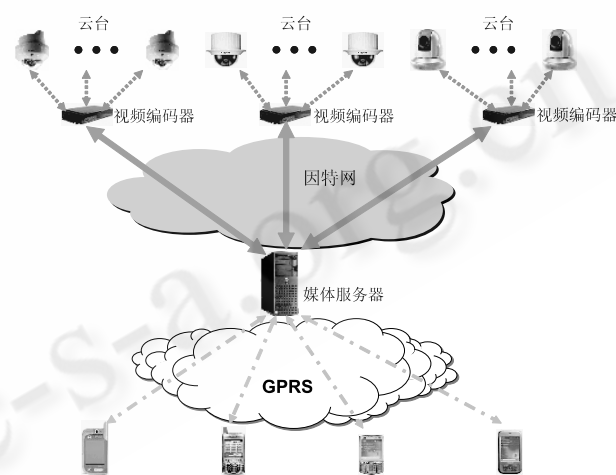


图 1 “手机视频监控系统”组网结构图

手机视频监控系统的组网结构图如图 1 所示。手机终端通过 GPRS 或其他手段(如 WAP、3G 等)接入媒体服务器,从媒体服务器获取视频数据并通过媒体服务器对云台进行 PTZ 控制。多个网络摄像机拍摄到的视频图像可以传送到一台视频编码器进行编码,从而提高了编码器的使用效率,降低了建设成本。编码器编码、流化后得到的视频数据流通过因特网传送到统一的媒体服务器平台,在用户数量巨大的情况下,可以

^① 本文通信作者

配置多台服务器,以实现流量均衡及系统容错。

在手机视频监控系统中组网结构图中,媒体服务器需要实现与编码器的通信及信令交换,以完成对云台的控制和媒体数据的获取,同时需要实现与手机终端的通信,将来自编码器的媒体流发送到手机,并接收来自手机的控制信令;手机需要实现媒体的解码及播放功能,同时需要与媒体服务器进行通信,以实现媒体流的获取,并根据用户的操作情况,发送相应的信令给媒体服务器,并接受媒体服务器的监控数据流、云台控制的结果信息等。

3 手机监控通信协议的设计与实现

本文所设计与实现的手机视频监控通信协议(MMCP)仅为手机视频监控系统中智能手机和媒体服务器之间的控制信令,不涉及编码器与媒体服务器之间通信的细节。在手机监控通信协议中需要解决如下几个问题:

- (1) 用户登录与认证;
- (2) 视频媒体流的收发;
- (3) 摄像头及云台的 PTZ 控制

为简化信令协议设计,本文设计与实现的 MMCP 采用 TCP 收发相关的媒体数据及信令消息,默认 TCP 端口号为 8899,也可以预先对通信双方的端口号设为其它适当的值。

3.1 MMCP 信令格式

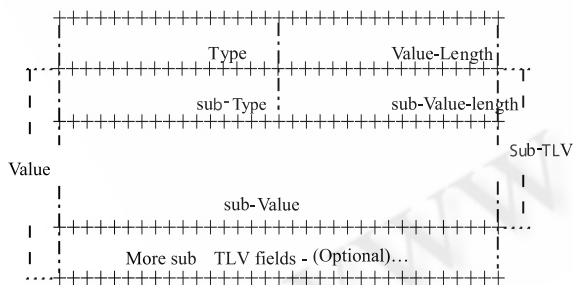


图 2 MMCP 信令格式

MMCP 的信令格式如图 2 所示。总体上,协议的所有报文均采用典型的 TLV 格式 (Type - Length - Value)。为了区分协议细节,其中的 Value 又包含了一个或者多个 TLV 类型的字段。分别描述如下:

(1) Type : 占一个字节的字段,其取值表明该 MMCP 报文的类型。各种 Type 字段可能的取值及其含义如表 1 所示。

表 1 MMCP 的报文类型

Type 取值	对应报文类型	描述
1	Login	登录请求
2	Logout	注销请求
3	Play	要求服务器发送视频的请求
4	Data	包含视频数据的数据报文
5	Response	某些请求报文对应的响应报文
6	Preset	请求设置预置位的报文
7	PTZ	请求对远程摄像头进行 PTZ 控制
8	Keep Alive	用于维护手机和媒体服务器之间连接的保活报文
9	Preset Info	用于向手机发布预置位信息

(2) Value - Length : 占一个字节的字段,说明报文中 Value 字段的长度,单位为字节。Length 字段的长度不包含 Type 字段及 Length 字段本身的长度。

(3) Value : 构成整个 TLV 报文中的 Value 字段,该 Value 字段又按照类似的格式划分为各个子类型、子类型报文的长度及子类型报文的的内容字段。Value 字段的长度由 Value - Length 给出。以下将一组 Sub - Type、Sub - Value - Length 和 Sub - Value 字段简称作为一个 sub - TLV 字段。

不同 Type 取值的 MMCP 报文,至少包含一个 sub - TLV 字段,也可能包含多个 sub - TLV 字段。对于只有一个 sub - TLV 字段的报文(如 Logout 报文),其中的 sub - Type 和 sub - Length 字段实际上是多余的,但为了方便协议实现,同时考虑到将来可能的协议扩展,仍然采用图 2 所示的报文结构。

在 sub - TLV 字段中的各个字段说明如下:

(1) sub - Type : 占 1 字节,说明 sub - Value 字段的类型。

(2) sub - Value - Length : 占 1 字节,说明 sub - Value 字段以字节为单位的长度。

(3) sub - Value : 包含了 sub - TLV 字段的协议内容。

MMCP 协议报文按照二进制流处理,不需要进行字节对齐。

3.2 登录请求报文

只有经过授权的用户才能利用手机查看远程监控画面。因此用户必须先进行登录,登录时须提供合法的用户名和密码,指明所登录的媒体服务器的域名或

者 IP 地址,以及希望查看的监控摄像头编号。

手机端软件在获取上述信息之后,向指定的媒体服务器发送登录请求报文(Login Request Packet),登录请求报文中包含了明文形式的用户名以及经过 MD5 加密的登录口令。

登录请求报文格式如图 2 所示,其中 Value 字段包含 2 个 sub-TLV 字段。当 sub-Type 字段等于 1 时,对应的 sub-Value 字段为 UTF 编码的用户名;当 sub-Type 字段等于 2 时,对应的 sub-Value 字段为 MD5 加密的登录口令。

3.3 播放请求报文

用户成功登录后,手机端软件向媒体服务器发送播放请求(Play Request),要求服务器将监控视频的编码流发到手机。

播放请求报文格式如图 2 所示,其中 Value 字段包含 1 个 sub-TLV 字段。sub-Type 字段等于 1,sub-Value 字段为 4 个字节的摄像头编号(即 sub-Value-Length 字段等于 4)。

3.4 注销请求报文

用户关闭手机视频监控软件后,手机端自动向媒体服务器发送注销请求(Logout Request),通知服务器该手机已经停止查看监控画面。

注销请求报文采用图 2 中的格式,但其中不包含任何 sub-TLV 字段。Type 字段取值为 2(见表 1),Value-Length 字段取值为 0。

3.5 保活报文

为了避免手机端故障导致的失效连接仍然占用系统资源,手机端监控软件在成功登录之后,每隔 30 秒需要向媒体服务器发送一个保活报文(Keep Alive Packet)。

保活报文采用图 2 中的格式,但其中不包含任何 sub-TLV 字段。Type 字段取值为 8(见表 1),Value-Length 字段取值为 0。

3.6 云台控制

本文设计的 MMCP 协议可以在手机端通过媒体服务器控制远程摄像头云台的动作,包括:左转、右转、上仰、下俯,统称为 PTZ(Pan-Tilt-Zoom)控制。同时还可以进行设置预置位及转到特定预置位的操作。以实现按需查看被监控环境的目标。

3.6.1 PTZ 控制

手机端向媒体服务器发送 PTZ 控制(PTZ Control)报文,通过媒体服务器控制监控摄像头的 PTZ 动作。

PTZ 控制报文采用图 2 中的格式,Value 字段包含 1 个 sub-TLV 字段。其中,sub-Type 字段等于 1,sub-Value-Length 字段固定为 4,sub-Value 字段为 4 个字节,其取值及含义如表 2 所示:

表 2 PTZ Control 报文中 sub-Value 的取值及含义

sub-Value 取值	含义
21	云台上仰
22	云台下俯
23	云台左转
24	云台右转
25	云台自动左右扫描

3.6.2 预置位控制

预置位控制包括设置预置位、转到预置位及清除预置位三种操作。为了便于管理,每个预置位除了有一个编号之外,还有一个预置位名称。预置位的编号采用 4 个字节。

手机端向媒体服务器发送预置位控制报文(Preset Point Control Packet),并通过服务器实现对摄像头的预置位控制。为了避免多个手机终端对预置位进行控制导致的冲突,需要在媒体服务器端对手机终端进行授权。

预置位控制报文的 Type 字段等于 6(见表 1),报文采用图 2 中的格式。这类报文的 Value 字段包含两个 sub-TLV 字段。

当用于设置预置位时,两个 sub-TLV 字段分别包含需设置的预置位的编号(sub-Type 等于 1)及预置位的名称(sub-Type 等于 2)。预置位名称采用 UTF 编码。

当用于控制摄像头转到预置位或者用于清除已有的预置位时,第一个 sub-TLV 字段包含需转到或者清除的预置位的编号,其 sub-Type 等于 1,sub-Value 字段即为预置位编号(Index);第二个 sub-TLV 字段则用于表明预置位控制操作的类型,其 sub-Value-Length 字段等于 0,即不需要 sub-Value 字段,若 sub-Type 等于 3 则表示转到预置位,若 sub-Type 等于 4 则表示清除预置位。

3.7 服务器端报文

3.7.1 响应报文

服务器在接收到来自手机端的各种控制报文之后,需要将操作结果反馈给手机端。为此使用了一组响应报文(Response Packet)。服务器响应报文的 Type 字段等于 5(见表 1),报文采用图 2 中的格式。

这类报文的 Value 字段包含一个或两个 sub - TLV 字段。第一个 sub - Type 字段等于 0,则表示对操作成功的响应,等于 1 则表示对操作失败的响应,sub - Value 字段为 4 字节的响应码。响应码的具体取值及含义见表 3。

表 3 响应码的取值及含义

响应码	含义
1	云台上仰操作的结果
2	云台下俯操作的结果
3	云台左转操作的结果
4	云台右转操作的结果
5	云台自动左右扫描操作的结果
6	设置预置位操作的结果
7	清除预置位操作的结果
8	转到预置位操作的结果

如果响应报文包含第二个 sub - TLV 字段,则其 sub - Type 字段的取值见表 4, sub - Value 字段为与响应相关的信息。sub - Type 字段的取值及含义如表 4 所示。

表 4 响应报文中,第二个 sub - Value 的取值及含义

sub - Type 取值	含义
2	sub - Value 包含端口号
3	sub - Value 包含预置位编号

例如,如果设置编号为 1 的预置位成功,则响应报文的 Type 字段等于 5(表示这是一个响应报文),该报文包含两个 sub - TLV 字段,第一个 sub - Type 等于 0(表示操作成功),第一个 sub - Value 字段等于 6(表示这是“设置预置位”操作的结果,见表 3);第二个 sub - Type 等于 3(表示后面的 sub - Value 包含预置位编号),第二个 sub - Value 字段等于 1(表示操作对象是编号为 1 的预置位)。

3.7.2 预置位信息

媒体服务器需要保存已经设置的预置位信息,在手机成功登录服务器之后,服务器必须将与手机登录时指定的摄像头相关的预置位信息发送到手机端。

服务器使用预置位信息报文(Preset Point Info Packet)通知手机端已经设置的预置位信息,包括预置位编号及预置位名称。

预置信息报文的 Type 字段等于 9(见表 1),报文

采用图 2 中的格式。这类报文的 Value 字段包含多组 sub - TLV 字段,每组有两个 sub - TLV 字段,对应一个已经设置的预置位。每个 sub - TLV 字段中的 sub - Type 字段,如果等于 1,表示其对应的 sub - Value 字段为预置位的 4 字节的编号,如果 sub - Type 等于 2,则表示其对应的 sub - Value 字段为预置位的用 UTF 编码的名称。

3.7.3 媒体数据报文

媒体服务器在接收到成功登录的手机端发送过来的播放请求之后,如果有足够的系统资源,就向手机终端发送视频数据流。来自媒体服务器的视频数据封装在媒体数据报文中,使用手机登录时建立的连接向手机端发送数据(即带内信令)。

媒体数据报文的 Type 字段等于 4(见表 1),采用图 2 中的格式,其中 Value 字段包含来自编码器的媒体数据,而不是 sub - TLV 字段。

4 协议实现及效果

视频编码采用 H. 264 格式,图像分辨率为 QCIF,帧率 3FPS。根据以上设计,在手机端采用 J2ME^[1],在服务器端采用 C + +^[2],实现了 MMCP 协议。可以灵活方便的实现多个手机登录服务器以及一台服务器同时控制多个摄像头的功能,在查看视频图像的同时,手机端可以控制摄像头进行 PTZ 动作及预置位控制操作。

经测试,利用 J2ME 开发的手机端软件可以在 Nokia N72、多普达 P800 及 Palm Treo 650 等手机上运行,测试时视频图像的显示延迟低于 7 秒,画面比较清晰、流畅,效果令人满意。

参考文献

- Jonathan Knudsen, Sing Li. Beginning J2ME: From Novice to Professional. Third Edition, Apress, 2005. 100 - 356.
- Stanley B. Lippman, José Lajoie, Barbara E. Moo. C + + Primer. Fourth Edition, Addison - Wesley Professional, 2005. 300 - 412.
- S60 3rd Edition SDK Supporting Feature Pack 2, for MIDP, User's Guide, Version 0.9, Nokia, 2007.